广电监管数据资产管理系统的关键技术及实现

摘　要：广播电视监管工作中的相关数据已经累积到相当多的程度，传统的存储和数据处理技术体系的建设维护越来越不能适应数据量增长速度和业务上的需求，在保证原有系统平稳运行的前提下，非常需要利用新的技术体系对数据进行迁移、扩容和备份，以 Hadoop 为代表的开源大数据处理和存储软件框架给我们提供了一套可行的方案。本文依据数据资产管理系统的实际建设经验，对其中的关键技术和建设过程进行分析和介绍。

关键词：大数据；Hadoop；HDFS；MapReduce；MPP；Kettle；数据仓库
中图分类号：TN948.6　　文献标识码：A
文章编号：1671-0134（2019）07-111-03　　DOI：10.19483/j.cnki.11-4653/n.2019.07.036

文 / 王洋

前言

广播电视监管事业经历了从模拟到数字、从地面到卫星、从单机到网络等一系列技术体系的重大变革，随着技术的不断发展，广播电视节目的频率数、频道数、播出时长也在快速增加，开路电视、有线电视、卫星电视、IPTV 等多种传输方式也在快速出现，因此，监管工作所面临的数据量在急剧膨胀，早期工作中是以结构化的业务数据为主，主要是广播电视节目在播出和传输过程中的相关技术性指标，广播电视内容以实时监听监看为主，录音录像保存时间很有限。在过去的十几年中，大规模海量数据的存储和处理方案在商业公司和开源软件社区中涌现出来，这源于互联网的发展导致互联网公司迫切需要发展适合于大数据环境下的全新技术体系，达到低成本、易扩展、并发性高、兼容不同类型数据的目的。现实的应用需求使得一些技术方案快速成熟，这样就为广播电视监管工作提供了可行的解决方案，大量音视频节目的存储成为可能。与此同时，海量数据处理技术和分析技术的发展也能够满足不断提升的监管业务需求。

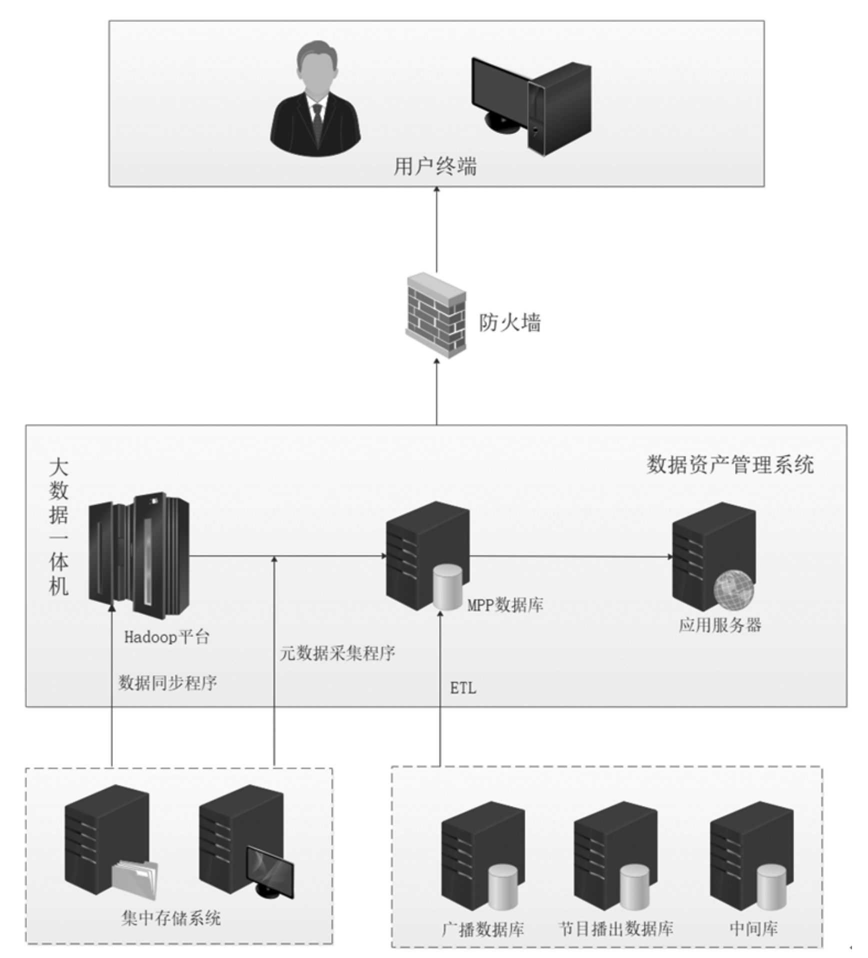


图 1　数据资产管理系统技术架构图

就目前来讲，广电监管工作主要用到两大类数据，一类是对应广播电视节目的音视频文件，另一类是相关监管业务数据，如广播效果情况、电视节目播出表等，前者为非结构化数据，后者为结构化数据。音视频文件的存储容量已经达到 PB 级，每天的增量达到了 TB 级，包括了大部分的卫星和有线广播电视节目，由于存储容量已达极限，需要定期从存储中删除旧的音视频文件，业务数据虽然容量只是 GB 级，但某些数据增长快速，已经达到亿级的条目数。随着以上两类数据的积累，对数据存储的扩展性和安全性要求逐渐提升，必须要在数据读取性能和可靠性之间取得平衡，需要以较低成本较简单的操作实现较高的并发访问和分布式存储。此外，海量数据带来数据分析的需求，需要新的技术方案弥补现有系统数据仓库能力不足的问题。最后，类似 Oracle 这样的传统商业型数据库虽然能很好地满足关系型数据管理和分析的需求，但使用成本较高，分布式架构操作复杂，源代码不开放，存在数据资产的安全性问题。综上所述，我们需要针对音视频文件和相关业务数据组成的数据资产管理需求，利用开源软件社区的成熟大数据技术方案弥补已有技术体系的不足，实现性价比更高、更加安全的数据资产管理系统。

1. 技术架构介绍

各个监管业务系统之间彼此封闭，每个系统内形成封闭的回路，没有达到有效的数据共享，因此，首先需要对数据进行有效整合，消除数据孤岛，建立统一规范的数据标准，形成对数据的全景统计，在此基础上开发可视化数据统计分析功能，实现对数据资产的准确掌握和有效管理，辅助建立跨系统、跨部门、跨地区的沟通协调机制。下图是根据上述需求形成的技术架构图，目前已经完全实现并正常运行。

下面一层是原有的数据存储系统，包括音视频数据和业务数据，中间一层是以大数据一体机为形式的新系统，包括用于存放音视频文件的 Hadoop 平台和存放文件元数据和业务数据的 MPP 数据库，基于这些数据建立数据资产管理系统，包括数据资产目录、数据存量统计、数据增量统计、数据流量监控、作业监控等功能，这些功能经过防火墙在上面一层呈现在用户终端上。

集中存储系统采用的是浪潮的 Lustre 系统，存放着历年的广播电视节目音视频文件，广播数据库是指广播播出的技术性业务数据，节目播出数据库是指广播电视节目编目数据，中间库是其他系统同步过来的技术性业务数据，这三个数据库均采用 Oracle 一体机，以 RAC(Real Application Cluster) 方式进行存储。以上两类数据经过大数据一体机中的 Hadoop 软件和 ETL 工具，实现了数据转移和提取，形成数据资产管理系统可以直接利用的数据源。

2. 关键技术及实现

根据技术架构的要求，我们接下来要决定采用哪些

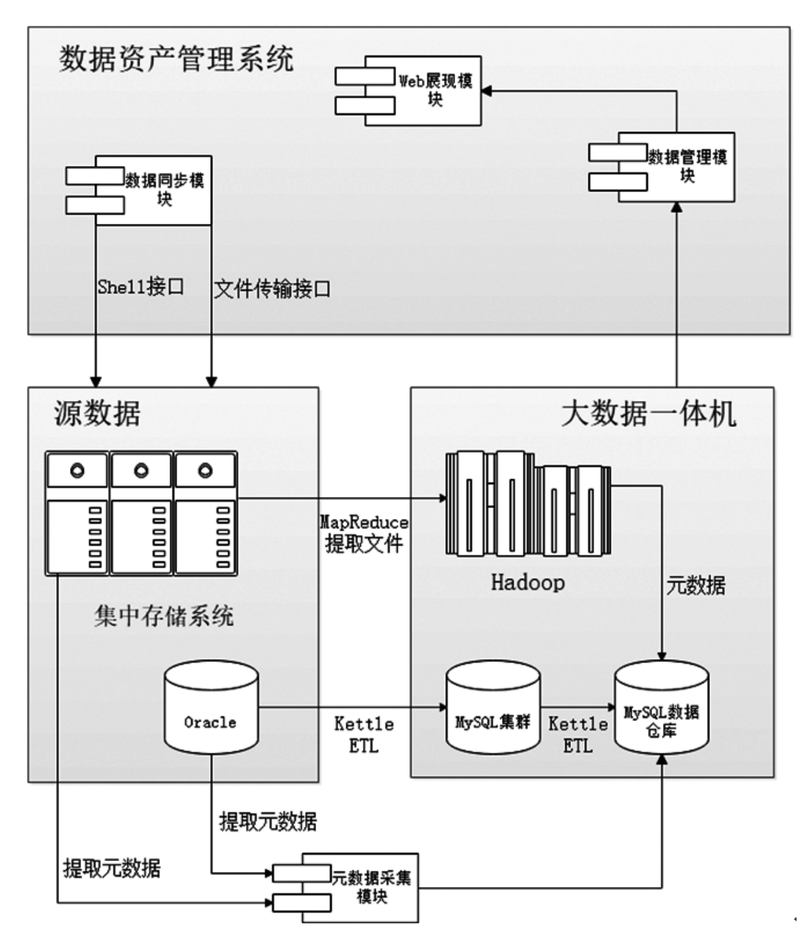


图 2 数据资产管理系统软件结构图

关键技术实现其中的数据同步、ETL 等功能。下图体现了不容软件模块之间的数据流向和用到的一些技术或产品。

集中存储系统中的 Lustre 系统是一种基于 Linux 的开源分布式文件系统，Lustre 的最大特点是读写能力很强，特别是对于大文件，利用 Stripe 条带化技术实现文件的并行访问，元数据服务器存储文件的元数据信息，对象存储服务器用于存储文件，Lustre 的主要问题是对数据的保护比较差，底层的对象存储结构不支持数据的备份和回复，另外，元数据存在单点故障的问题，采用了共享存储层支持主备切换，一旦共享存储层出现问题导致元数据无法访问，进而影响整个系统。我们选择了 Hadoop 2.0 来读取和存放音视频文件，HDFS 支持多个副本在不同节点的存储，可以实现数据的保护，并具有数据校验的功能，文件的操作虽然不像 Lustre 支持 POSIX 接口，只能通过 API 实现，读取性能不如 Lustre，但适合需要长期保存的文件以及建立数据仓库。另外，HDFS 生态环境非常丰富，有很多强大的开源框架可以使用，这一点要比 Lustre 好很多。在这个项目中我们使用了 2 个管理节点和 20 个数据节点，管理节点使用 Zookeeper 实现平台集群的负载均衡，数据节点通过 HDFS 存放了数百 TB 的音视频文件。

集中存储系统文件信息读取的接口采用 MapReduce 实现，MapReduce 程序对集中存储系统的内容进行解析，提取出集中存储系统的音视频文件的元数据，并完成两项任务，一是将音视频文件同步到大数据 Hadoop 平台上；二是获取到音视频文件的属性信息并同步到 MPP 数据库中。数据同步管理策略包括文件同步和元数据同步两个方面，元数据同步每 15 分钟进行一次，主要是通过 find 命令查询当天日期目录下最近一个小时新增的文件，并将元数据保存到数据库中。同步定时任务通过 Linux 的 crontab 命令实现，元数据抽取的日志文件位于与 jar 包同级的 logs 文件夹里。文件同步每天晚上 0 点进行，根据数据库中前一天的元数据信息，将前一天所有的文件复制到 HDFS 中，在主要文件完成之后根据文件增加情况和同步需求不定期执行文件同步。

对于结构化数据的存储我们采用了基于 MysSQL 的 MPP（Massively Parallel Processor）数据库，MPP 是大规模并行处理的意思，系统由很多松耦合处理单元组成，每个单元内的 CPU 都有自己的私有资源，在每个单元内都有操作系统和管理数据库的实例副本。MPP 架构数据库具有的特征是任务并行执行、数据分布式存储、分布式计算、私有资源、横向扩展和 Share Nothing 架构。横向扩展是 MPP 数据库的主要设计目标，MPP 数据库支持严格的关系模型，比如 SQL92、加扩展、加存储过程，支持事务、保证数据强一致性，所解决的问题包括提升数据处理性能和数据处理量。与各个节点使用自己私有资源的 Shared Nothing 架构不同，原有数据库采用的 Oracle RAC 属于 Shared Disk 架构，各个节点使用自己的 CPU 和内存，磁盘存储共享，也就是数据共享，当存储性能达到瓶颈时，增加节点便不能获得并行能力的扩展，另外，Oracle RAC 源代码不开放，存放于 Oracle 一体机中，这种一体机价格较为昂贵。

MPP 数据库集群共有 6 个节点，共分为三类，即管理节点、数据节点和 SQL 节点。管理节点有 2 个，作用是管理集群内的其他节点，如提供配置数据、启动并停止节点、运行备份等，由于这类节点负责管理其他节点的配置，应在启动其他节点之前首先启动这类节点，管理节点用命令 ndb_mgmd 启动。数据节点有 4 个，用于保存集群中的数据，数据节点的数目与副本的数目相关，用命令 ndbc 启动。SQL 节点有 4 个，用于访问集群数据，通常 SQL 节点使用命令 mysqld –ndbcluster 启动，或将 ndbcluster 添加到 my.cnf 后使用 mysqld 启动。管理节点负责管理集群配置文件和集群日志，集群中的每个节点从管理节点检索配置数据，并请求确定管理节点所在位置，数据节点内出现新的事件时，将关于这类时间的信息传输到管理节点，然后将这类信息写入集群日志。

从 Oracle RAC 到 MPP 数据库的 ETL 过程采用了开源 ETL 工具 Kettle，ETL 即数据抽取（Extract）、转换（Transform）、装载（Load）的过程，它是构建数据仓库的重要环节。传统的 ETL 工具有集中执行、对服务器性能要求高等缺点，针对这些缺点，本项目采用了一种基于分布式原理的 ETL 技术，该系统在分布式文件系统基础上实现了集群分布式 ETL 流程。该分布式 ETL 系统具有较高的可扩展性和吞吐效率，同时能够自动实现负载均衡，执行效率高。分布式 ETL 技术包括四个组件：Spoon、Pan、Chef、Kithcen。Spoon 实现了通过图形界面来设计 ETL 转换过程，通过 SPoon 设计 ETL 工作的转换和作业，转换定义了 ETL 抽取地源、目标和抽取规则，是 ETL 的主体，作业控制着转换的执行，Pan 实现了批量运行由 Spoon 设计的 ETL 转换；Pan 是一个后台执行的程序，没有图形界面，Chef 可以创建任务，任务通过允许每个转换、任务、脚本等，更有利于自动化更新数据仓库的复杂工作，任务将会被检查，看看是否正确地运行了；Kitchen 允许你批量使用 CHEF 设计的任务，例如使用一个时间调度器，Kitchen 也是一个后台运行的程序。分布式 ETL 集群共有 8 个节点，2 个主节点和 6 个从节点。

结语

以上我们从业务发展的需要出发，介绍了数据资产管理系统的技术架构和关键技术，以及一些具体的实现方式，经过这个项目，音视频和业务数据资源得到了整合，数据存储容量更大并且更加安全，而且易于扩展，为进一步的数据挖掘工作奠定了良好的基础。随着大数据开源软件社区新的成熟项目不断出现，数据资产管理系统可以充分融合新的工具实现功能的不断增加。媒

（作者单位：国家广播电视总局监测数据处理中心）